



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/054,366	01/22/2002	James W. Cooper	YOR9-2001-0647US1	8382
29683	7590	06/02/2005	EXAMINER	
HARRINGTON & SMITH, LLP 4 RESEARCH DRIVE SHELTON, CT 06484-6212			BETIT, JACOB F	
		ART UNIT		PAPER NUMBER
		2164		

DATE MAILED: 06/02/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	10/054,366	COOPER ET AL.
	Examiner	Art Unit
	Jacob F. Betit	2164

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 27 October 2004.
- 2a) This action is FINAL. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1-33 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) Claim(s) 8,9,23 and 24 is/are allowed.
- 6) Claim(s) 1-7,10-22 and 25-33 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.



SAM RIMELL
PRIMARY EXAMINER

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____ | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| | 6) <input type="checkbox"/> Other: _____ |



DETAILED ACTION

Remarks

1. In response to communications filed on 21- October-2004, claims 1, 2, 4, 6-11, 15-17, 19, 21-26, and 31-33 are amended per applicant's request. Claims 1-33 are presently pending in the application.

Claim Rejections - 35 USC § 102

2. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

3. Claims 1-3, 7, 16-18, 22, and 31 are rejected under 35 U.S.C. 102(e) as being anticipated by Gomes et al. (U.S. patent No. 6,615,209 B1).

As to claim 1, Gomes et al. teaches a method for processing data representing documents, comprising:

for individual documents of a set of documents, executing a software program to obtain a list of salient terms found in each document (see column 10, line 42 through column 12, line 58);
comparing the list of salient terms for a first document to the list of salient terms for a second document (see column 12, line 59 through column 13, line 41); and

declaring the first document to be substantially identical to, or substantially similar to, the second document if some predetermined number of salient terms are found in each of the lists of the first document and the second document (see column 8, lines 37-60).

As to claims 2 and 17, Gomes et al. does not teach wherein if the predetermined number is about 90% of the salient terms or greater the first document is declared to be substantially identical to the second document (see column 12, line 59 through column 13, line 41).

As to claims 3 and 18, Gomes et al. teaches wherein the set of documents is obtained in response to a search query made to a data communications network (see column 5, lines 43-65).

As to claims 7 and 22, Gomes et al. teaches wherein the step of executing a software program assigns to each salient term a collection-level importance ranking or Information Quotient (IQ), and wherein the IQ is considered during the step of comparing (see column 13, lines 1-22).

As to claim 16, Gomes et al. teaches a system for processing data representing documents comprising, for individual documents of a set of documents, a processor for executing a software program to obtain a list of salient terms found in each document (see column 10, line 42 through column 12, line 58) and for comparing the list of salient terms for a first document to the list of salient terms for a second document (see column 12, line 59 through column 13, line 41), said processor being operable for declaring the first document to be substantially identical to, or

Art Unit: 2164

substantially similar to, the second document if some predetermined number of salient terms are found in each of the lists of the first document and the second document (see column 8, lines 37-60).

As to claim 31, Gomes et al. teaches a computer program recorded on a computer-readable media, said computer program comprising instructions for directing a data processor to process data representing documents by, for individual documents of a set of documents, obtaining a list of salient terms found in each document (see column 10, line 42 through column 12, line 58); comparing the list of salient terms for a first document to the list of salient terms for a second document (see column 12, line 59 through column 13, line 41); and declaring the first document to be substantially identical to, or substantially similar to, the second document if some predetermined number of salient terms are found in each of the lists of the first document and the second document (see column 8, lines 37-60).

4. Claims 10, 12-13, 25, 27-28, and 32 are rejected under 35 U.S.C. 102(e) as being anticipated by Pugh et al. (U.S. patent No. 6,658,423 B1).

As to claim 10, Pugh et al. teaches method for processing data representing documents, comprising:

for individual ones of documents, executing a software program to obtain a list of salient terms found in each document (see column 11, lines 1-60);

computing a document signature for each document from the list of salient terms obtained for the document (see column 11, line 61 through column 13, line 67, where the number of lists is one);

comparing the document signature for a first document to the document signature for a second document (see column 14, lines 1-42); and

declaring the first document to be substantially identical to the second document if the document signatures are substantially equal (see column 14, lines 25-35).

As to claims 12 and 27, Pugh et al. teaches wherein the documents are obtained in response to a search query made to a data communications network, and where the steps of comparing and declaring are executed in substantially real time as the documents are returned by the query (see column 19, lines 7-9).

As to claims 13 and 28, Pugh et al. teaches wherein the documents are obtained in response to a search query made to a data communications network, where the steps of comparing and declaring are executed in substantially real time as the documents are received from the data communications network, and for a case where a received document is found to be substantially identical to an already received document, returning only one of the documents in response to the search query (see column 19, lines 7-9).

As to claim 25, Pugh et al. teaches a system for processing data representing documents, comprising, for individual documents of a set of documents, a processor for executing a software

program to obtain a list of salient terms found in each document (see column 11, lines 1-60), for computing a document signature for each document from the list of salient terms obtained for the document (see column 11, line 61 through column 13, line 67, where the number of lists is one); for comparing the document signature for a first document to the document signature for a second document (see column 14, lines 1-42); and for declaring the first document to be substantially identical to the second document if the document signatures are equal (see column 14, lines 25-35).

As to claim 32, Pugh et al. teaches a computer program recorded on a computer-readable media, said computer program comprising instructions for directing a data processor to process data representing documents by, for individual ones of documents, obtaining a list of salient terms found in each document (see column 11, lines 1-60); computing a document signature for each document from the list of salient terms obtained for the document (see column 11, line 61 through column 13, line 67, where the number of lists is one); comparing the document signature for a first document to the document signature for a second document (see column 14, lines 1-42); and declaring the first document to be substantially identical to the second document if the document signatures are equal (see column 14, lines 25-35).

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

- (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person

Art Unit: 2164

having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 4-6 and 19-21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gomes et al. (U.S. patent No. 6,615,209 B1) in view of Kathrow et al. (U.S. patent No. 6,263,348 B1).

As to claims 4 and 19, Gomes et al. does not teach further comprising storing the lists of salient terms in a database.

Kathrow et al. teaches further comprising storing the lists of salient terms in a database (see column 5, lines 26-43).

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Gomes et al. to include further comprising storing the lists of salient terms in a database.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Gomes et al. by the teachings of Kathrow et al. because further comprising storing the lists of salient terms in a database would allow the current invention to operate periodically (see Kathrow et al., column 5, lines 29-35).

As to claims 5 and 20, Gomes et al. does not teach further comprising computing a signature for each document, and storing the computed document signature.

Kathrow et al. teaches further comprising computing a signature for each document (see column 5, lines 11-25), and storing the computed document signature (see column 5, lines 26-43).

Art Unit: 2164

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Gomes et al. to include further comprising computing a signature for each document, and storing the computed document signature.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Gomes et al. by the teachings of Kathrow et al. because further comprising computing a signature for each document, and storing the computed document signature would allow both similar and identical files to be found at any periodic time (see Kathrow et al., abstract, and see column 5, lines 29-35).

As to claims 6 and 21, Gomes et al. does not teach further comprising computing a signature for each document, and storing the computed document signature in association with the list of salient terms for each document.

Kathrow et al. teaches further comprising computing a signature for each document (see column 5, lines 11-25), and storing the computed document signature in association with the list of salient terms for each document (see column 5, lines 26-43).

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Gomes et al. to include further comprising computing a signature for each document, and storing the computed document signature in association with the list of salient terms for each document.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Gomes et al. by the teachings of Kathrow et al. because further comprising computing a signature for each document, and storing the computed

Art Unit: 2164

document signature in association with the list of salient terms for each document would allow both similar and identical files to be found at any periodic time (see Kathrow et al., abstract, and see column 5, lines 29-35).

7. Claims 11, 26, and 33 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pugh et al. (U.S. patent No. 6,658,423 B1) in view of Piosenka et al. (U.S. patent No. 4,993,068).

As to claims 11, 26, and 33, Pugh et al. does not teach wherein the step of computing a document signature computes a hash code for each term of the list of salient terms, and then sums all of the hash codes to form the document signature.

Piosenka et al. teaches wherein the step of computing a document signature computes a hash code for each term of the list of salient terms, and then sums all of the hash codes to form the document signature (see column 7, lines 7-30).

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Pugh et al. to include wherein the step of computing a document signature computes a hash code for each term of the list of salient terms, and then sums all of the hash codes to form the document signature.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Pugh et al. by the teachings of Piosenka et al. because wherein the step of computing a document signature computes a hash code for each term of the list of salient terms, and then sums all of the hash codes to form the document signature would result in high probability that digital signatures of modified blocks would differ from the signatures of blocks that are the same (see Piosenka et al., column 7, lines 7-30).

8. Claims 14-15 and 29-30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Pugh et al. (U.S. patent No. 6,658,423 B1) in view of Kathrow et al. (U.S. patent No. 6,263,348 B1).

As to claims 14 and 29, Pugh et al. does not teach further comprising storing the computed document signatures in a database.

Kathrow et al. teaches further comprising storing the computed document signatures in a database (see column 5, lines 26-43).

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Pugh et al. to include further comprising storing the computed document signatures in a database.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Pugh et al. by the teachings of Kathrow et al. because further comprising storing the computed document signatures in a database would allow the current invention to operate periodically (see Kathrow et al., column 5, lines 29-35).

As to claim 15, Pugh et al. does not teach further comprising storing the computed document signature in association with the list of salient terms for each document.

Kathrow et al. teaches further comprising storing the computed document signature in association with the list of salient terms for each document (see column 5, lines 26-43).

Art Unit: 2164

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Pugh et al. to include further comprising storing the computed document signature in association with the list of salient terms for each document.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Pugh et al. by the teachings of Kathrow et al. because further comprising storing the computed document signature in association with the list of salient terms for each document would allow the current invention to operate periodically (see Kathrow et al., column 5, lines 29-35).

As to claim 30, Pugh et al. as modified, teaches further comprising storing the computed document signature in association with the list of terms for each document (see Kathrow et al., column 5, lines 26-43).

Allowable Subject Matter

9. Claims 8-9 and 23-24 are allowed.

Response to Arguments

10. Applicant's arguments filed 27-October-2004 have been fully considered but they are not persuasive.

In response to the applicant's arguments that "the Gomes patent at the very least fails to teach or suggest obtaining a list of salient terms for each document and then utilizing these list of

salient terms for comparing the similarity of one document to another”, the arguments have been fully considered but are not deemed persuasive. Marriam-Webster online dictionary defines salient as standing out conspicuously, prominent, of notable significance. By using “query-relevant information” Gnomes et al. uses terms that are salient, prominent or of notable significance, rather than the entire document, to determine how similar one document is to another. Since this information is obtained from the query that is input into the system, Gnomes et al. teaches obtaining this information. The applicant’s arguments state that a “salient term” is defined as “a single word or a multi-word term that meets a predetermined confidence”, and that a “salient term” is not “required to be based even in part on query relevant information”. Neither of these requirements is made in the claims pending in the application. Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

In response to the applicant’s arguments that “the Pugh patent reference at the very least fails to teach or suggest ‘...computing a document signature for each document from the list of salient terms obtained for the document’”, the arguments have been fully considered but are not deemed persuasive. Pugh et al. teaches removing “short or common words or terms” so that they are not processed during the extraction operation. This leaves the salient (prominent or of notable significance) terms to be processed during the extraction procedure. This list or more prominent terms is obtained by extracting them from the original document.

Conclusion

11. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

12. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jacob F. Betit whose telephone number is (571) 272-4075. The examiner can normally be reached on Monday through Friday 9 am to 5 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dov Popovici can be reached on (571) 272-4083. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

jfb
18 May 2005



SAM RIMELL
PRIMARY EXAMINER